TRACKING MULTIPLE OBJECTS USING INTENSITY-GVF SNAKES

Jonas De Vylder[1,*] , *Daniel Ochoa*[1,2], *Wilfried Philips*[1], *Laury Chaerle*[3], *Dominique Van Der Straeten*[3]

[1]Department of Telecommunications and Information Processing, IBBT - Image Processing and Interpretation Group, Ghent University, St-Pietersnieuwstraat 41, B-9000 Ghent, Belgium
[2]Facultad de Ingeniera en Electricidad y Computacin, Escuela Superior Politcnica del Litoral (ESPOL) Campus Gustavo Galindo, Km. 30. 5 via Perimetral, Apartado 09-01-5863, Guayaquil, Ecuador
[3]Department of Physiology, Unit Plant Hormone Signalling and Bio-imaging, Ghent University, Ledeganckstraat 35, B-9000 Ghent, Belgium

ABSTRACT

Active contours or snakes are widely used for segmentation and tracking. Multiple object tracking remains a difficult task, characterised by a trade off between increasing the capturing range of edges of the object of interest, and decreasing the capturing range of other edges. We propose a new external force field which is calculated for every object independently. This new force field uses prior knowledge about the intensity of the object of interest. Using this extra information, this new force field helps in discriminating between edges of interest and other objects. For this new force field, the expected intensity of an object must be estimated. We propose a technique which calculates this estimation out of the image.

1. INTRODUCTION

The reliable estimation of objects' features in images is a time consuming task. It demands skilled technicians who spend time identifying and measuring objects of interest in the image. Although software drawing tools can ease this work, when the measurements have to be monitored over time this approach becomes impractical. The problem can be complicated further if the objects of interest changes their location and shape. Since we are interested in measuring each object individually, the logical approach is to apply techniques that on the one hand provide accurate segmentation, with minimal human intervention, and on the other reduce the complexity of the tracking process. Region based segmentation techniques [1, 2, 3] require every frame to be segmented and the resulting segments linked to those found in the next frame. The disadvantage of this approach is that in case of errors, for instance over-segmentation, the number of segments in each frame might be different. In that case, complex splitting and merging operation have to be implemented to match objects correctly. Model based approaches [4, 5, 6, 7] , that incorporate motion and shape information into the segmentation process, provide a more solid framework when a suitable representation of the object of interest is available. We choose the active contour framework for this work because it allows the integration of region and contour constraints and avoids the problems of the afore mentioned techniques. The number of objects in every frame remains constant and the segmented contours in one frame can be used as initial contour in the next. For objects in isolation this initial contour will converge towards the object's new location and deform to match the object's new shape.

However when objects with different contrast lay close to each other, the contour corresponding to the object with lower contrast will be attracted to the object with higher contrast. This behavior is due to the nature of active contour formulation, that drives the contour toward regions with high contrast. To cope with this problem, we propose a new external force field, i.e. Intensity Gradient Vector Flow field. This force field is calculated for every object separately. It combines edge strength and expected intensity of the object of interest, therefore increasing the possibilities that the corresponding contour will converge correctly.

This paper is arranged as follows. The next section provides a detailed description of active contours. Both the classical and gradient vector flow snakes are explained. In section 3 our proposed algorithm is presented. Section 4 shows an example of tracking leafs in a sequence of thermal images. This example is compared to other snake formulations. Finally, results are discussed in section 5.

2. ACTIVE CONTOURS

2.1 Snakes

The classical snake model proposed by Kass et al. [4], defines the active contour as a parametric curve, $\mathbf{r}(s) = (x(s), y(s))$, that moves in the spatial domain until the energy functional in Eq. 1 reaches its minimum value.

$$E_{snake} = \int \Big(E_{int}(\mathbf{r}(s)) + E_{ext}(\mathbf{r}(s)) \Big) ds. \tag{1}$$

E_{int} and E_{ext} represent the internal and external energy, respectively. The internal energy enforces smoothness along the contour. A common internal energy function is defined as follows:

$$E_{int}(\mathbf{r}(s)) = \Big(\alpha \mid \mathbf{r}'(s) \mid^2 + \beta \mid \mathbf{r}''(s) \mid^2 \Big) / 2 \tag{2}$$

where α and β are weighting parameters, $\mathbf{r}'$ and $\mathbf{r}''$ are the fist and second derivative of $\mathbf{r}(s)$ with respect to s. The first term, also known as tension energy, prevents the snake to remain attracted to isolated points. The second term, known as bending energy, prevents the contour of developing sharp angles. Constraints based on more complex shape models, such as Fourier descriptors, have also been reported in literature [8, 9].

The external energy is derived from the image, so that the snake will be attracted to features of interest. Given a gray level image $I(x,y)$, a common external energy is defined as:

* Corresponding author: phone: +32 9 264 3416, email: jonas.deylder@telin.ugent.be

$$E_{ext}(I(x,y)) = -\mid \nabla I(x,y) \mid^2 \tag{3a}$$

or

$$E_{ext}(I(x,y)) = -\left|\nabla\big(G_\sigma(x,y) * I(x,y)\big)\right|^2 \tag{3b}$$

where ∇ is the gradient operator, $G_\sigma(x,y)$ a 2D Gaussian kernel with standard deviation σ and where $*$ is the convolution operator.

Minimizing the energy function of eq. 1 results in solving the following associated Euler-Lagrange equations:

$$\alpha \frac{d^2x(s)}{ds^2} - \beta \frac{d^4x(s)}{ds^4} - \frac{\partial E_{ext}}{\partial x} = 0 \tag{4a}$$

$$\alpha \frac{d^2y(s)}{ds^2} - \beta \frac{d^4y(s)}{ds^4} - \frac{\partial E_{ext}}{\partial y} = 0 \tag{4b}$$

This can be seen as a force balance equation:

$$F_{int} + F_{ext} = 0 \tag{5}$$

These equations can be solved using gradient descent by treating $\mathbf{r}(s)$ as a function of time, i.e. $\mathbf{r}(s,t)$. The partial derivative of $\mathbf{r}$ with respect to t is then

$$\frac{dx(s,t)}{dt} = \alpha \frac{d^2x(s,t)}{ds^2} - \beta \frac{d^4x(s,t)}{ds^4} - \frac{\partial E_{ext}}{\partial x} \tag{6a}$$

$$\frac{dy(s,t)}{dt} = \alpha \frac{d^2y(s,t)}{ds^2} - \beta \frac{d^4y(s,t)}{ds^4} - \frac{\partial E_{ext}}{\partial y} \tag{6b}$$

When the snake stabilizes, i.e. when an optimum is found, the terms $\frac{dx(s,t)}{dt}$ and $\frac{dy(s,t)}{dt}$ vanish.

2.2 Gradient Vector Flow

The external force field defined in the previous section requires a good initialization, close to the object boundary, in order to segment the object. This limitation is caused by the nature of the external force field, whose vectors point towards the object only in the proximity of the object's boundary. As we move away from the boundary the external fields rapidly become zero, therefore reducing the possibilities that a contour located in such regions will converge correctly. To overcome this problem, Xu and Prince [10] proposed another external force field $\mathbf{v}(x,y) = (u(x,y), v(x,y))$. This vector field minimizes the following energy functional:

$$E_{GVF}(u,v) = \iint \mu\Big(\frac{du}{dx}^2 + \frac{du}{dy}^2 + \frac{dv}{dx}^2 + \frac{dv}{dy}^2\Big) + \mid \nabla f \mid^2 \mid \mathbf{v} - \nabla f \mid^2 \, dxdy \tag{7}$$

where μ is a nonnegative parameter expressing the degree of smoothness of the field $\mathbf{v}$ and where f is an edge map, e.g. $\mid \nabla I \mid$. The first term of Eq.7 keeps the field $\mathbf{v}$ smooth, whereas the second term forces the field $\mathbf{v}$ to resemble the original edge force in the neighborhood of edges. This new external force is called *gradient vector flow* (GVF) field. The GVF-field can be found by solving the following associated Euler-Lagrange equations:

$$\mu \nabla^2 u - \Big(u - \frac{\partial f}{\partial x}\Big)\Big(\frac{\partial f}{\partial x}^2 + \frac{\partial f}{\partial y}^2\Big) = 0 \tag{8a}$$

$$\mu \nabla^2 v - \Big(u - \frac{\partial f}{\partial y}\Big)\Big(\frac{\partial f}{\partial x}^2 + \frac{\partial f}{\partial y}^2\Big) = 0 \tag{8b}$$

where ∇^2 is the Laplacian operator.

3. INTENSITY GRADIENT VECTOR FLOW

The GVF field formulation effectively increases the range an object can attract a contour, consequently relaxing the initialization constrains. However, when an image contains multiple objects the GVF field can bias the contour evolution towards the objects with higher contrast. We illustrate this problem in Fig. 1 where two objects, a high contrast square and a low contrast circle, their gradient and the image GVF field are depicted. As can be seen in Fig. 1.b the square displays the stronger gradient. Also note in Fig. 1.c that in the region located between the objects, the resulting GVF field points toward the square object. This occurs because the GVF field equations assign non-zero values to regions with no gradient data. If a contour is initialized at a small distance from the circle, the segmentation can yield unpredictable results such as in Fig. 1.d. A simple strategy to solve this problem is increasing the weighting parameters in Eq.2. This will reduce the search space of the contour and the possibilities that it gets trapped under the influence of neighboring objects. The negative aspect of this approach is that it also reduces the possible shapes that can be segmented.

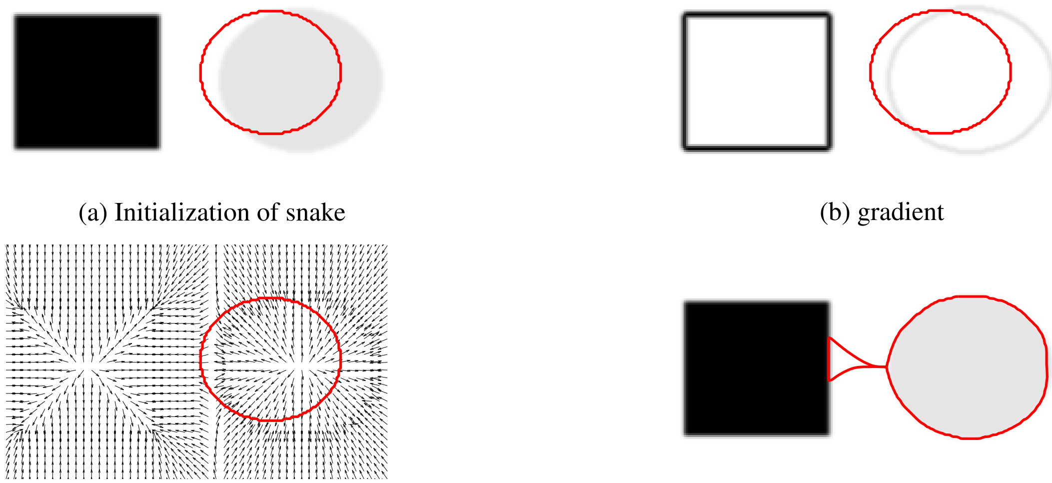
Another possible solution is limiting the influence of objects far away. This strategy however needs a good initialization. This is not always feasible in tracking applications, where the resulting contour of frame t is used as an initialization of the contour of frame $t+1$. So the GFV-force should range as far as the objects of interest can move between two subsequent frames. Another disadvantage of tracking is the risk of error propagation, namely that wrong detected segments will be used as initialization in the next frame.

To cope with this problem two schemes have been reported: one focused on solving the multiple object problem and the other focused on solving the motion problem. In the first approach, Tang [11] and Cohen [5] propose to manipulate the external force field so that gradients towards a certain direction are favored. This approach requires a good estimation of object's location. In the second approach, the snake tries to anticipate the new location of the object [9, 12, 13], therefore a better initialization is assumed and the influence of other objects limited. These algorithms utilize some form of prior knowledge about the motion, in practice this knowledge is not always available which limits the number of applications.

The solution we propose belongs to the first group, but instead of manipulating the force field based on location, we assume to have prior knowledge about the intensity of the object. In the remaining of this section we will first define the operators necessary for our technique. Then our proposed technique itself will be explained. We conclude with a method to estimate the expected intensity.

3.1 Neighboring intensity

In order to incorporate object features into the segmentation process we evaluate the object intensity values. We decided to use intensity information because it is more reliable than edge data to characterize the objects with a fairly homogeneous interior. A problem in this approach is the estimation



(a) Initialization of snake (b) gradient

(c) GVF-force field (d) segmentation using a GVF-snake

Figure 1: Example of segmenting the circle using a GVF-snake.

of intensity values at the object's boundary. Different objects can have similar values, particularly in very noisy images. We propose the neighboring intensity operator Ψ to overcome such problem:

$$\Psi_b(x,y,I) = I\left(x + b\frac{\partial I(x,y)}{\partial x}, y + b\frac{\partial I(x,y)}{\partial y}\right) \tag{9}$$

It samples the intensity of pixels located at distance b from the x.y position. The sign of b indicates if the measure is taken in the direction of the gradient or its opposite.

3.2 Intensity Model

We assume that intensity of an object Z can be modeled using a Normal probability distribution i.e. $P(Z) \sim N(\mu, \sigma^2)$ with a constant σ^2 for all frames. The σ^2 value can be estimated during initialization, while the μ values is updated in every frame. The ability of the contour to discriminate between objects using intensity information, depends on these models. If the μ of different objects are too close to each other, or the σ^2 is too large, the proposed method might not work.

3.3 Intensity Gradient Vector Flow

Once the intensity model parameters are estimated, we can assign to each pixel the probability that it belongs to a given object. Since the boundary pixels are of major importance, but as mentioned before they do not provide a reliable measure of the object intensity, we estimate the corresponding probability using the output of the neighboring intensity operator instead.

$$\Gamma(x,y,I,E[I]) = \max\begin{cases} P(z = \Psi_b(x,y,I) \mid \mu = E[I]) \\ P(z = \Psi_{-b}(x,y,I) \mid \mu = E[I])) \end{cases} \tag{10}$$

Where $E[I]$ is the expected intensity of the object of interest. So the mean of the objects intensity model is defined as $E[I]$. Using the neighboring intensity operator, we look in both directions of the gradient. This is because we consider the possibility that the object's surroundings can be darker or lighter than the object itself. The parameter b is set to ensure that the neighboring intensity operator samples pixels of the objects interior. The Ψ operator looks for example for a distance of two pixels in the direction of the gradient. Instead of choosing b constant, it might be interesting to make b variable for each pixel, i.e. let b be defined in function of the gradient strength.

This new intensity map can be combined with the GVF method. In eq.7 we redefine the edge map for every object, i.e.

$$f_{E[I]} = \mid \nabla I \mid \Gamma(x,y,I,E[I]) \tag{11}$$

This new external force is called *intensity gradient vector flow* (IGVF) field.

3.4 Intensity Estimation

The expected intensity of an object could be predicted out of the result of previous frames. An other possibility is estimating it based on the image and the initialization of the snake. If the initialization is not too far from the true object, we can estimate the intensity based on the intensity of the initialization. To discard the influence of background or other objects, we can use the median or a weighted mean as an estimation of the expected intensity. If the object initialization area is more outside the object than inside, these techniques will result in wrong estimations.

4. RESULTS

To test our method we have segmented a sequence of images which monitored a sugar beet seedling plant. This time-lapse sequence was captured with a thermal camera. These thermal images are grayscale images, such as can be seen in Fig.2.a. Every hour a new image was taken. In this sequence the 4 leafs of the plant all move in different directions, at different speeds. Since this movement does not seem to have a clear motion model, we can not incorporate any prior knowledge about the motion in our tracking methods. There is however an intensity difference between the two big leafs and the two small leafs.

As an initialization for the snakes in the first frame, a watershed algorithm [1] was used. In Fig.2.b an example of this initialization is shown. In figure c the gradient of the thermal image is shown. As can be seen, prevails the gradient of the two big leafs the image, specially near the stalk of the 'yellow leaf'. Finally, Fig.2.d shows the Γ image of the leaves, as expected intensity, the median of the yellow contour in Fig.2.b was chosen. Here the small leaves prevail the image.

In Fig.3 the segmentation results are shown. In Fig.3.a-b the results of classical snakes are shown. In the first example the strong attraction of the big leafs is visible. Even though the error is still small, the yellow contour is already attracted away from its true contour towards the big leaf. Fig.3.b does not suffer from this problem because it puts high constraints on the shapes of the snake. It achieves this by defining high values for α and β in Eq.2. In Fig.3.c the intensity-edge map is shown. The contours of all four leaves have similar edge strengths. This results in a good segmentation as can be seen in Fig.3.d.

In order to quantitatively evaluate the proposed technique, a full sequence was manually segmented and compared to the segmentation results for both the proposed and the classical active contours. The provided ground truth consists of 54 images, all containing four leaves. The Dice coefficient is used as a similarity measure such as is done in [14]. Given $S_{\mathbf{r}}, S_{\mathbf{true}}$, the regions defined by respectively $\mathbf{r}$ and the true contour, the Dice coefficient is defined as:

$$DC(S_{\mathbf{r}}, S_{\mathbf{true}}) = \frac{2Area(S_{\mathbf{r}} \cap S_{\mathbf{true}})}{Area(S_{\mathbf{r}}) + Area(S_{\mathbf{true}})} \quad (12)$$

The Dice coefficient returns 1 if the segments are identical, and 0 if they are totally different. The proposed method results in an average Dice coefficient of 0.89 with an average improvement of 0.05 compared to the classical snake with the same parameters.

This improvement might not seem spectacular, but note that every error in a frame has the chance of propagating to other frames, if the active contour framework is used for tracking. Fig.4 shows the tracking results for frame 7 of the same sequence as Fig.3. Both GVF-snakes result in wrong segments. In Fig.4.a the error of Fig.3.a has propagated. Where as in Fig.4.b the consequence of the strong constraint on the shape is visible. As can be seen in Fig.4.d, the IGVF-snake does not have problems with multiple objects, nor with the shape of the objects.

5. DISCUSSION AND CONCLUSION

In this paper a new variant on GVF-field , *intensity gradient vector flow* field, is defined. This new force field takes both gradient and the expected intensity in consideration. The IGVF-field is useful for tracking multiple objects, where the objects of interest have different intensities. The power to discriminate between different objects using IGVF, depends on how different the object's intensities are. The expected intensity of an object can be estimated out of the initialization. If however the initialization of the object is too far, there exist a risk of moving towards wrong objects. This error mainly occurs when the object initializations area is more outside the object than inside, what results in a wrong estimation of the intensity. To omit this problem the expected intensity can be predicted out of previous frames. Some future work could involve the testing of the proposed method for the segmentation of objects with a constant texture. The IGVF-field for this application could be build on the response of a texture filter instead of the intensity image. We believe this new technique can work in a wide range of applications where the classical active contours have failed.

6. ACKNOWLEDGMENT

The authors are grateful to Erik De Bruyne, SESVanderhave NV, for providing the plant material. Jonas De Vylder is funded by the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT). The authors would also like to thank the reviewers for there interesting input.

REFERENCES

[1] J. De Bock and W. Philips, "Line segment based watershed segmentation," *Advances in Computer Vision and Computer Graphics, Third International Conference on Computer Vision/Computer Graphics*, pp. 579–586, 2007.

[2] J. De Bock, R. Pires, P. De Smet, and W. Philips, "A fast dynamic border linking algorithm for region merging," *Advanced Concepts for Intelligent Vision Systems, ACIVS, 8th international conference*, pp. 232–241, 2006.

[3] M. Vincent, "Morphological grayscale reconstruction in image analysis: application and efficient algorithms," *IEEE transactions on image processing*, vol. 2, pp. 176–201, 1993.

[4] M. Kass, A Witkin, and D. Terzopoulos, "Snakes: active contour models," *International journal of computer vision*, pp. 321–331, 1988.

[5] L.D. Cohen, "On active contour models and balloons," *CVGIP: Image Understanding*, vol. 53, pp. 211–218, 1991.

[6] T. Cootes and C. Taylor, "Statistical models of appearance for medical image analysis and computer vision," *Proceedings SPIE Medical Imaging*, 2001.

[7] D. Freedman, R. Radke, T. Zhang, m. Jeong, Y. Lovelock, and G. Chen, "Model-based segmentation of medical imagery by matching distributions," *IEEE transactions on medical imaging*, vol. 24, pp. 281–292, 2005.

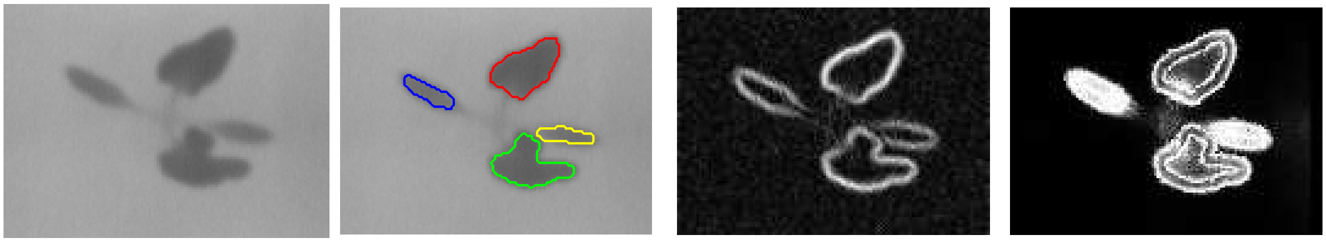
[8] M.-A. Charmi, S. Derrode, and S. Ghorbel, "Fourier-based geometric shape prior for snakes," *Pattern Recognition Letters*, vol. 29, pp. 897–904, 2008.

[9] A.P. Goobic, M.E. Welser, S.T. Acton, and K. Ley, "Biomedical application of target tracking in clutter," *Proc. 35th Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 88–92, 2001.

[10] C. Xu and J.L. Prince, "Snakes, shapes and gradient vector flow," *IEEE transactions on image processing*, vol. 7, pp. 359–369, 1998.

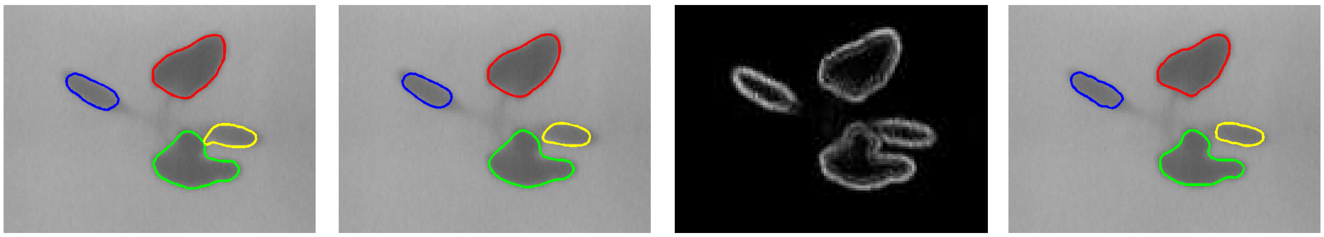
[11] J. Tang, "A multi-direction gvf snake for the segmentation of skin cancer images," *Pattern Recognition*, 2008, doi:10.1016/j.patcog.2008.09.007.

[12] N. Ray and S. Acton, "Motion gradient vector flow: An external force for tracking rolling leukocytes with shape



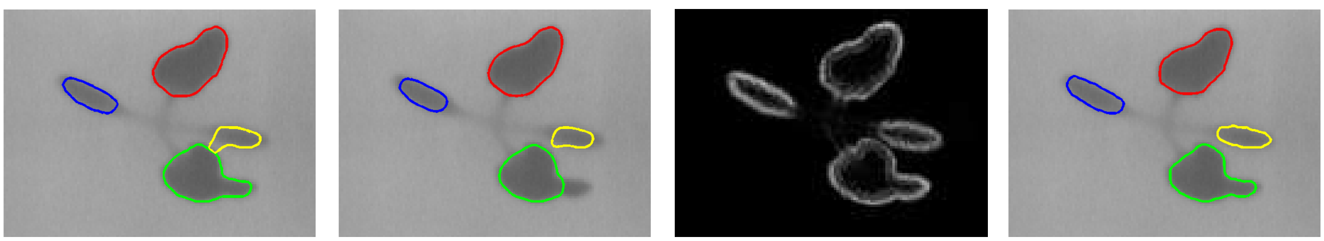
(a) An example of a thermal image | (b) Initialization using watershed | (c) The gradient of Fig. a | (d) The Γ image derived from yellow region

Figure 2: Frame 1 of a thermal time lapse sequence, and some of its preprocessing steps.



(a) GVF-snake, with low internal energy | (b) GVF-snake, with high internal energy | (c) intensity edge map see eq.(11) | (d) IGVF-snake, with low internal energy

Figure 3: Example of segmentation of frame 1 using different active contours



(A) GVF-snake, with low internal energy | (b) GVF-snake, with high internal energy | (c) intensity edge map see eq.(11) | (d) IGVF-snake, with low internal energy

Figure 4: Example of tracking, results for frame 7 using different active contours

and size constrained active contours," *IEEE TRANSACTIONS ON MEDICAL IMAGING*, vol. 23, pp. 1466–1478, 2004.

[13] G. Tsechpenakis, K. Rapantizikos, N. Tsapatsoulis, and S. Kollias, "A snake model for object tracking in natural sequences," *Signal processing: image communication*, vol. 19, pp. 219–238, 2004.

[14] O. Bernard, O. Friboulet, P. Thevenaz, and M. Unser, "Variational b-spline level-set: A linear filtering approach for fast deformable model evolution," *IEEE transactions on image processing*, vol. 18, pp. 1179–1191, 2009.